Two kinds of Phase transitions in a Voting model

M Hisakado* and S Mori†

February 27, 2013

*Standard & Poor's, Marunouchi 1-6-5, Chiyoda-ku, Tokyo 100-0005, Japan

†Department of Physics, School of Science, Kitasato University, Kitasato 1-15-1, Sagamihara, Kanagawa 252-0373, Japan

Abstract

In this paper, we discuss a voting model with two candidates, C_0 and C_1. We consider two types of voters–herders and independents. The voting of independents is based on their fundamental values; on the other hand, the voting of herders is based on the number of previous votes. We can identify two kinds of phase transitions. One is an information cascade transition similar to a phase transition seen in Ising model. The other is a transition of super and normal diffusions. These phase transitions coexist. We compared our results to the conclusions of experiments and identified the phase transitions in the upper limit of the time t by using analysis of human behavior obtained from experiments.

*[1] masato_hisakado@standardandpoors.com
†[2] mori@sci.kitasato-u.ac.jp

1 Introduction

In general, collective herding behavior poses interesting problems in several cross fields such as sociology [1], social psychology [2], ethnology [3][4], and economics. Statistical physics offers effective tools to analyze these phenomena caused by collective herding behaviors, and the associated field is known as sociophysics [5]. For example, in statistical physics, anomalous fluctuations in financial markets [6][7] and opinion dynamics [8][9][10] have been related to percolation and the random field Ising model.

To estimate public perception, people observe the actions of other individuals; then, they make a choice similar to that of others. Because it is usually sensible to do what other people are doing, collective herding behavior is assumed to be the result of a rational choice. This approach can sometimes lead to arbitrary or even erroneous decisions as a macro phenomenon. This phenomenon is known as an information cascade [11]

In our previous paper, we introduced a sequential voting model that is similar to a Keynesian beauty contest [12][13][14]. At each time step t, one voter votes for one of two candidates. As public perception, the tth voter can see all previous votes, i.e. $(t-1)$ votes. There are two types of voters–herders and independents–and two candidates. Herders are also known as copycat voters; they vote for each candidate with probabilities that are proportional to the candidates' votes. We refer to these herders as analog herders. We investigated a case wherein all the voters were herders [15]. In such a case, the voting process is a Pólya process, and the voting rate converges to a beta distribution in a large time limit of t [16].

Next, we added independents to the analog herders [13]. In the upper limit of t, the independents cause the distribution of votes to converge to a Dirac measure against herders. This model contains three phases–two super diffusion phases and a normal diffusion phase. We refer to the transition in this model as a transition of super and normal diffusions. These transitions can be seen in several fields [17]. If herders constitute the majority or even half of the total voters, the voting rate converges to a Dirac measure slower than in a binomial distribution. These two phases have different speeds of convergence that are slower than in a binomial distribution. If independents constitute the majority of the voters, the voting rate converges at the same rate as that in a binomial distribution. If the independents vote for the correct candidate rather than for the wrong candidate, the model does not include the case wherein the majority of the voters choose the wrong can-

didate. The herders affect only the speed of the convergence; they do not affect the voting rates for the correct candidate.

Next, we consider herders who always choose the candidate with a majority of the previous votes, which is visible to them [18]. We refer to these herders as digital herders. Digital herders exhibit stronger herd behavior than analog herders. We obtained exact solutions when the voters comprised a mix of digital herders and independents. As the fraction of herders increases, the model features a phase transition beyond which a state where most voters make the correct choice coexists with one where most of them are wrong. This phase transition is referred to as information cascade transition.

Here, we discuss a voting model with two candidates, C_0 and C_1. We set two types of voters–independents and herders. The voting of independents is based on their fundamental values. They collect information independently. On the other hand, the voting of herders is based on the number of previous votes, which is visible to them. In this study, we consider the case wherein a voter can see all previous votes.

From experiments, we observed that human beings exhibit a behavior between that of digital and analog herders. We obtained the probability that a herder makes a choice under the influence of his/her prior voters' votes. The probability can be fitted by a tanh function. If the difference between numbers of voters for candidates C_0 and C_1 is small, the probability that a herder chooses the candidate receiving a majority of the previous votes increases rapidly. If the difference between numbers of voters for candidates C_0 and C_1 is large, the probability becomes constant. In this paper, we discuss rich phases of the models in which herders exhibit behavior that can be fitted to a tanh function. We identify two types of phase transitions–information cascade transition and transition between super and normal diffusions. Furthermore, we discuss the phases of models that we obtained from experiments [19].

The remainder of this paper is organized as follows. In section 2, we introduce our voting model and mathematically define the two types of voters–independents and herders. In section 3, we derive a stochastic differential equation. In section 4, we discuss information cascade transition by using the stochastic differential equation. In section 5, we discuss the phase transition between normal and super diffusions and we demonstrate the coexistence of these phase transitions. In section 6, we verify these transitions through numerical simulations. In section 7, we discuss social experiments from the viewpoint of our models. Finally, the conclusions are presented in section 8.

2 Model

We model the voting of two candidates, C_0 and C_1; at time t, C_0 and C_1 have $c_0(t)$ and $c_1(t)$ votes, respectively. In each time step, one voter votes for one candidate; the voting is sequential. Hence, at time t, the tth voter votes, after which the total number of votes is t. Voters are allowed to see all the previous votes for each candidate; thus, they are aware of public perception.[1]

There are two types of voters–independents and herders; we assume an infinite number of voters. The independents vote for C_0 and C_1 with probabilities $1-q$ and q, respectively. Their votes are independent of others' votes, i.e. their votes are based on their fundamental values.

Here, we set C_0 as the wrong candidate and C_1 as the correct candidate in order to validate the performance of the herders. We can set $q \geq 0.5$ because we believe that independents vote for the correct candidate C_1 rather than for the wrong candidate C_0. In other words, we assume that the intelligence of the independents is virtually accurate.

On the other hand, the herders' votes are based on the number of previous votes. We use the following functions. If the numbers of votes are $c_0(t)$ and $c_1(t)$ at time t, a herder votes for C_1 at time $t+1$ with the following probability:

$$q_h = \frac{1}{2}[\tanh \lambda \{ \frac{c_1(t)}{(c_0(t)+c_1(t))} - \frac{1}{2} \} + 1]. \tag{1}$$

If C_1 receives the majority votes, i.e. $c_1(t)/(c_0(t)+c_1(t)) > 1/2$, the ratio of votes for C_1, denoted by q_h, increases to 1 exponentially. λ reflects how the previous answers affect a voter's choice. If λ is positive and large, the voter has high confidence in the previous votes. We use function (1) for our social experiments, and it fits well to human behaviors in the experiments [19]. In our experiments, we can estimate $\lambda = 3.80$. If $c_0(t) = c_1(t)$, herders vote for C_0 and C_1 with the same probability, i.e. $1/2$. This model can also be derived from Bayes' theorem (see Appendix A). In this case, λ is not constant and it increases as the number of votes increases, according to the central limit theorem. The case is on the assumption that all voters are independents. Hereafter, we treat λ as a parameter.(see figure 1)

In the upper limit of λ, herders vote for the candidate with the majority votes. If $c_0(t) > c_1(t)$, the herders vote for candidate C_0, whereas if $c_0(t) < c_1(t)$, they vote for candidate C_1. These herders are known as digital herders

[1] The tth voter can see $c_0(t-1)$ and $c_1(t-1)$ votes at time t.

[18]. We expand (1) as follows:

$$q_h \sim \frac{1}{2} + \frac{\lambda}{2}(\frac{c_1(t)}{(c_0(t)+c_1(t))} - \frac{1}{2}) - \frac{\lambda^3}{6}(\frac{c_1(t)}{(c_0(t)+c_1(t))} - \frac{1}{2})^3 + \ldots. \quad (2)$$

If $c_1(t)/(c_0(t)+c_1(t)) \sim 1/2$ and $\lambda = 2$, a herder votes for C_1 with a probability of $c_1(t)/(c_0(t)+c_1(t))$. These herders are known as analog herders with $q = 1/2$ [13]. Therefore, the herders who vote with probability (1) are a hybrid of analog and digital herders.

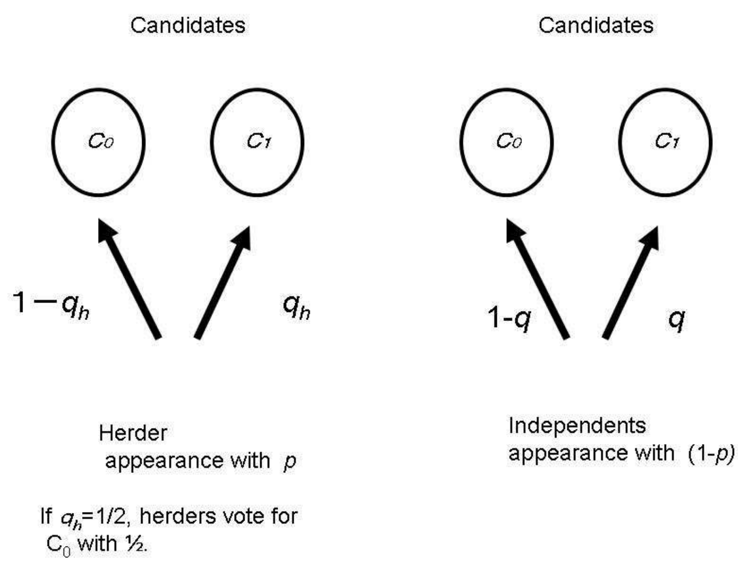


Figure 1: Representation of our model. Independents vote for C_0 and C_1 with probabilities $1-q$ and q, respectively. Herders vote for C_0 and C_1 with probabilities $1-q_h$ and q_h, respectively. We set the ratio of independents to herders as $(1-p)/p$. In the limit $\lambda \to \infty$, a herder is a digital herder. In the limit $\lambda = 2$, a herder is similar to an analog herder. In general, the herder is a hybrid between an analog herder and a digital herder with the parameter λ.

The independents and herders appear randomly and vote. We set the ratio of independents to herders as $(1-p)/p$. In this study, we mainly focus on the upper limit of t. This refers to the voting of infinite voters.

The evolution equation for candidate C_1 is

$$\begin{aligned} P(k,t) &= pq_h P(k-1,t-1) + p(1-q_h)P(k,t-1) \\ &\quad +(1-p)qP(k-1,t-1) + (1-p)(1-q)P(k,t-1). \end{aligned} \tag{3}$$

Here, $P(k,t)$ is the distribution of the number of votes k at time t for candidate C_1. The first and second terms of (3) denote the votes of the herders; the third and fourth terms denote the votes of the independents.

3 Stochastic Differential Equation

To investigate long-ranged correlations, we analyze in the limit $t \to \infty$. We can rewrite (3) as

$$c_1(t) = k \to k+1 : P_{k,t} = \frac{p}{2}[\tanh \lambda(\frac{k}{t-1} - \frac{1}{2}) + 1] + (1-p)q. \tag{4}$$

In the scaling limit $t = c_0(t) + c_1(t) \to \infty$, we define

$$\frac{c_1(t)}{t} \Longrightarrow Z. \tag{5}$$

Z is the ratio of voters who vote for C_1.

We define a new variable Δ_t such that

$$\Delta_t = 2c_1(t) - t = c_1(t) - c_0(t). \tag{6}$$

We change the notation from k to Δ_t for convenience. Then, we have $|\Delta_t| = |2k - t| < t$. Thus, Δ_t holds within $\{-t, t\}$. Given $\Delta_t = u$, we obtain a random walk model:

$$\begin{aligned} \Delta_t &= u \to u+1 : P_{\frac{u+t}{2},t} = \frac{p}{2}[\tanh\{\frac{\lambda u}{2(t-1)}\} + 1] + (1-p)q, \\ \Delta_t &= u \to u-1 : Q_{\frac{u+t}{2},t} = 1 - P_{\frac{u+t}{2},t}. \end{aligned}$$

We now consider the continuous limit $\epsilon \to 0$,

$$\begin{aligned} X_\tau &= \epsilon \Delta_{[t/\epsilon]}, \\ P(x,\tau) &= \epsilon P(\Delta_t/\epsilon, t/\epsilon), \end{aligned} \tag{7}$$

where $\tau = t/\epsilon$ and $x = \Delta_t/\epsilon$. Approaching the continuous limit, we can obtain the Fokker-Planck equation (see Appendix B):

$$\mathrm{d}X_\tau = [(1-p)(2q-1) + p\tanh(\frac{\lambda X_\tau}{2\tau})]\mathrm{d}\tau + \sqrt{\epsilon}. \tag{8}$$

Here, we change the variable X_τ to Y_τ by using the expression,

$$pY_\tau = X_\tau - (1-p)(2q-1)\tau. \tag{9}$$

We rewrite (8) by using Y_τ:

$$\mathrm{d}Y_\tau = \tanh\frac{p\lambda}{2\tau}(Y_\tau + \frac{(2q-1)(1-p)}{p}\tau)\mathrm{d}\tau + \sqrt{\epsilon}. \tag{10}$$

Using (5), (6), and (9), we can obtain the relations of the variables

$$2Z - 1 = \frac{X_\infty}{\tau} = \frac{pY_\infty}{\tau} + (1-p)(2q-1). \tag{11}$$

4 Information Cascade Transition

In this section, we discuss the information cascade transition. We observed this transition in the case of digital herders [18]. We are interested in the behavior in the limit $\tau \to \infty$. We consider the solution $Y_\infty \sim \tau^\alpha$, where $\alpha \leq 1$, since $\tanh x \leq 1$. The slow solution is $Y_\infty \sim \tau^\alpha$, where $\alpha < 1$ is hidden by the fast solution $\alpha = 1$ in the upper limit of τ. Hence, we can assume a stationary solution as

$$Y_\infty = \bar{v}\tau, \tag{12}$$

where $\bar{v}$ is constant. Substituting (12) into (10), we can obtain

$$\Delta Y_\infty = \tanh\frac{p\lambda}{2}(\bar{v} + \frac{(2q-1)(1-p)}{p})\Delta\tau = \bar{v}\Delta\tau. \tag{13}$$

The second equality is obtained from (12). Then, we obtain the equation

$$\bar{v} = \tanh\frac{p\lambda}{2}(\bar{v} + \frac{(2q-1)(1-p)}{p}). \tag{14}$$

This is the equation of state for Ising model. The second term on the RHS of (14) corresponds to the external field. In the cases of $q = 1/2$ and $p = 1$,

the external field disappears. If $\lambda > 2$, a phase transition occurs in the range $0 \leq p \leq 1$. As the number of herders increases, the model features a phase transition beyond which a state where most voters make the correct choice coexists with one where most of them are wrong. We refer to this transition as information cascade transition [18].

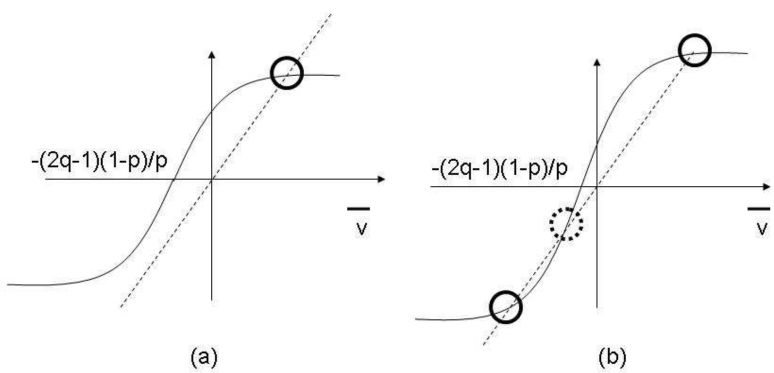


Figure 2: Solutions of self-consistent equation (14). It is the state equation of Ising model. (a) $p \leq p_c$ and (b) $p > p_c$. Below the critical point p_c, we can obtain one solution (a). We refer to this phase as the one peak phase. In contrast, above the critical point, we obtain three solutions. Two of them are stable and one is unstable (b). We refer to this phase as the two-peaks phase. In Ising model, the physical solution is the right side one only.

Equation (14) admits one solution below the critical point $p \leq p_c$ (see figure 2(a)) and three solutions for $p > p_c$ (see figure 2(b)). When $p \leq p_c$, we refer to the phase as the one peak phase. When $p > p_c$, the upper and lower solutions are stable; on the other hand, the intermediate solution is unstable. Then, the two stable solutions attain a good and bad equilibrium, and the distribution becomes the sum of the two Dirac measures. We refer to this phase as two-peaks phase. In Ising model with external fields, the

good equilibrium is only a physical solution. This is the difference between our voting model and Ising model. Here, we discuss two particular cases.

(1) Digital herder case: $\lambda = \infty$

In this case, the herders are considered to be digital herders. As shown in figure 2, the tanh function rises vertically at $\bar{v} = -\frac{(2q-1)(1-p)}{p}$. A phase transition occurs at $p_c = 1 - \frac{1}{2q}$. When $p \leq p_c$, we can obtain only one solution $\bar{v} = 1$. Using (11), we obtain the ratio of voters for C_1 as $Z = p + (1-p)q$. When $p > p_c$, we can obtain two stable solutions: $\bar{v} = 1$ and $\bar{v} = -1$. The solution $\bar{v} = -\frac{(2q-1)(1-p)}{p}$ is unstable. Using (11), the ratio of voters for C_1 is obtained as $Z = p + (1-p)q$ and $Z = q(1-p)$. $Z = p + (1-p)q$ shows a good equilibrium, whereas $Z = q(1-p)$ shows a bad equilibrium. This conclusion is consistent with the exact solutions [18].

(2) Symmetric independent voter case: $q = 1/2$

In this case, the external fields are absent. The self-consistent equation (14) becomes

$$\bar{v} = \tanh \frac{p\lambda}{2}(\bar{v}). \tag{15}$$

As shown in figure 2, the tanh function rises at $\bar{v} = 0$. If $\lambda \leq 2$, there is only one solution $\bar{v} = 0$ in all regions of p. In this case, Z has only one peak, at 0.5, which indicates the one peak phase. When herders are analog herders, we do not observe information cascade transition [18]. We observe only super and normal transitions (see section 5). On the other hand, if $\lambda > 2$, there are two stable solutions and an unstable solution $\bar{v} = 0$ above p_c. The votes ratio for C_1 attains a good or bad equilibrium. This is the so-called spontaneous symmetry breaking. In one sequence, Z is taken as $\bar{v}p/2 + 1/2$ in the case of a good equilibrium, or as $-\bar{v}p/2 + 1/2$ in the case of a bad equilibrium, where $\bar{v}$ is the solution of (15). This indicates the two-peaks phase, and the critical point is $p_c = 2/\lambda$.

5 Phase Transition of Super and Normal Diffusion Phases

In this section, we consider the phase transition of convergence. This type of transition has been studied when herders are analog [13]. The analog herders exhibit weaker herd behavior than digital herders. Depending on the convergence behavior, there are three phases. We expand Y_τ around the

solution $\bar{v}$.

$$Y_\tau = \bar{v}\tau + W_\tau. \tag{16}$$

Here, we set $Y_\tau \gg W_\tau$. This indicates $\tau \gg 1$. We rewrite (10) using (16) as follows:

$$\begin{aligned}
\mathrm{d}Y_\tau &= \bar{v}\mathrm{d}\tau + \mathrm{d}W_\tau \\
&= \tanh[\frac{p\lambda}{2\tau}(\bar{v}\tau + \frac{(2q-1)(1-p)}{p}\tau) + \frac{p\lambda}{2\tau}W_\tau]\mathrm{d}\tau + \sqrt{\epsilon} \\
&\sim \bar{v}\mathrm{d}\tau + \frac{p\lambda}{2\tau\cosh^2\frac{p\lambda}{2}(\bar{v} + \frac{(2q-1)(1-p)}{p})}W_\tau\mathrm{d}\tau + \sqrt{\epsilon} \\
&= \bar{v}\mathrm{d}\tau + \frac{p\lambda(1-\bar{v}^2)}{2\tau}W_\tau\mathrm{d}\tau + \sqrt{\epsilon}.
\end{aligned} \tag{17}$$

We use relation (14) and consider the first term of the expansion. Hence, we can obtain

$$\mathrm{d}W_\tau = \frac{p\lambda(1-\bar{v}^2)}{2\tau}W_\tau\mathrm{d}\tau + \sqrt{\epsilon}. \tag{18}$$

From Appendix C, we can obtain the phase transition of convergence. The critical point p_{vc} is the solution of

$$p_{vc} = \frac{1}{\lambda(1-\bar{v}^2)}, \tag{19}$$

and (14).

As shown in figure 2, at the critical point, the gradient of the tangent line at $\bar{v}$ is 1/2. If the gradient of the tangent line at $\bar{v}$ is under 1/2, the distribution converges as in a binomial distribution. We define γ as $Var(Z) = \tau^{-\gamma}$, where $Var(Z)$ is the variance of Z. The voting rate for C_1 converges as $Var(Z) = \tau^{-1}$: $\gamma = 1$. We refer to this phase as a normal diffusion phase. If the gradient of the tangent line at $\bar{v}$ is 1/2 or above 1/2, the voting rate converges at a speed slower than that in a binomial distribution. We refer to these phases as super diffusion phases. In one phase, the voting rate for C_1 converges to $\log(\tau)/\tau$, and in the other, the voting rate converges to $\gamma = p/p_{vc} - 2$.

(1) Digital herder case: $\lambda = \infty$

In this case, the herders are considered to be digital herders. In the one peak phase, where $p \leq p_c$, the only solution is $\bar{v} = 1$. Equation (18) represents the Brownian motion. The gradient of the tangent line at $\bar{v}$ is

0. Hence, the distribution converges as in a binomial distribution. In the two-peaks phase, where $p > p_c$, the solutions are $\bar{v} = \pm 1$. The gradient of the tangent line at $\bar{v}$ is 0. In each case, the distribution converges as in a binomial distribution. Hence, in all regions, the distribution converges as in a binomial distribution. In this limit, the phase transition of the convergence disappears and only information cascade transition is observed.

(2) Symmetric independent voter case: $q = 1/2$

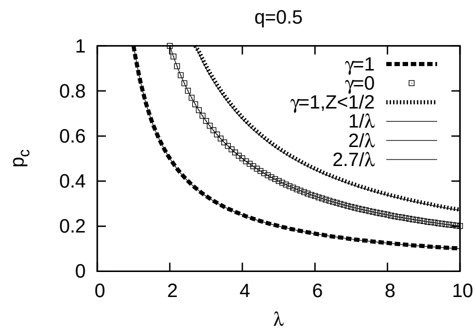
We consider the case $\lambda > 2$. In this case, we observe information cascade transition. If $\lambda \leq 2$, we do not observe information cascade transition and we can only observe a part of the phases, as described below.

In the one peak phase $p \leq p_c = 2/\lambda$, the only solution is $\bar{v} = 0$. p_c is the critical point of the information cascade transition. The first critical point of convergence is $p_{vc1} = 1/\lambda$. When $p \leq p_c$, p_{vc1} is the solution of (15) and (19). If $0 < p < p_{vc1}$, the voting rate for C_1 becomes 1/2, and the distribution converges as in a binomial distribution. If $p_c > p \geq p_{vc1}$, candidate C_1 gathers 1/2 of all the votes in the scaled distributions, too. However, the voting rate converges slower than in a binomial distribution. We refer to these phases as super diffusion phases. There are two phases, $p = p_{vc1}$ and $p_c > p > p_{vc1}$; these phases differ in terms of their convergence speed.

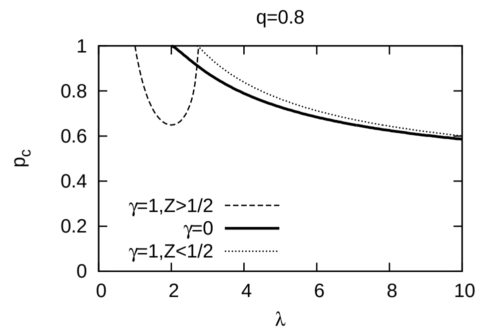
Above p_c, in the two-peaks phase, we can obtain two stable solutions that are not 0. At p_c, $\bar{v}$ moves from 0 to one of these two stable solutions. In one voting sequence, the votes converge to one of these stable solutions. If $p_c < p \leq p_{vc2}$, the voting rate for C_1 becomes $\bar{v}p/2 + 1/2$ or $-\bar{v}p/2 + 1/2$, and the convergence occurs at a rate slower than that in a binomial distribution. Here, $\bar{v}$ is the solution of (15). We refer to this phase as a super diffusion phase. p_{vc2} is the second critical point of convergence from the super to the normal diffusion phase, and it is the solution of the simultaneous equations (15) and (19) when $p > p_c$. We can estimate $p_{vc2} \sim \frac{2.5}{\lambda}$ by approximation.[2] In fact, with higher terms, we can estimate $p_{vc2} \sim \frac{2.7}{\lambda}$ (see figure 3(a)). On the other hand, if $p > p_{vc2}$, the voting rate for C_1 becomes $\bar{v}p/2 + 1/2$ or $-\bar{v}p/2 + 1/2$, too. But the distribution converges as in a binomial distribution. This is a normal diffusion phase. A total of six phases can be observed.

As discussed above, when $q = 0.5$, there is no difference between good and bad equilibriums. However, when $q = 0.8$, we observe a difference. $Z > 1/2$

[2] Using $\tanh x \simeq x - 1/3x^3$, equation (15), and the condition of critical point, we can obtain $p_{vc2} \simeq 2.5/\lambda$.



(a)



(b)

Figure 3: Phase diagram in space λ and p when (a) $q = 0.5$ and (b) $q = 0.8$. γ denotes the speed of convergence. $\gamma = 0$ is the boundary between the one peak phase and the two-peaks phase. The region below $\gamma = 0$ is the one peak phase. The region above $\gamma = 0$ is the two-peaks phase. $\gamma = 1$ is the between the normal diffusion phase and the super diffusion phase. In the one peak phase, the region below $\gamma = 1$ is the normal diffusion phase. In the two-peaks phase, two solutions are obtained. Each solution has a different speed of convergence, generally. When $q = 0.5$, there is no difference between the two solutions and the region below $\gamma = 1$ is the super diffusion phase. When $q = 0.8$, there is a difference between the good and the bad equilibriums. $Z > 1/2$ is the good equilibrium and $Z < 1/2$ is the bad equilibrium. The left side of $\gamma = 1$, $Z > 1/2$ in the two-peaks phase is the phase in which a good equilibrium is a super diffusion phase. The left side of $\gamma = 1$, $Z < 1/2$ is the phase in which a bad equilibrium is a super diffusion phase.

indicates the good equilibrium and $Z < 1/2$ indicates the bad equilibrium. Figure 3 shows the phase diagram in space λ and p when (a) $q = 0.5$ and (b) $q = 0.8$. γ is the speed of convergence. $\gamma = 0$ is the boundary between the one peak phase and the two-peaks phase. The region below $\gamma = 0$ is the one peak phase and that above $\gamma = 0$ is the two-peaks phase. $\gamma = 1$ is between the normal diffusion phase and the super diffusion phase. In the one peak phase, the region below $\gamma = 1$ is the normal diffusion phase and that above $\gamma = 1$ is the super diffusion phase. In the two-peaks phase, two solutions can be observed. When $q = 0.5$, we do not observe a difference between the two solutions and the region above $\gamma = 1$ is the normal diffusion phase. When

$q = 0.8$, we observe a difference in the speed of convergence between good and bad equilibriums. $Z > 1/2$ is the good equilibrium and $Z < 1/2$ is the bad equilibrium. The left side of $\gamma = 1$, $Z > 1/2$ in the two-peaks phase is the phase in which a good equilibrium is the super diffusion phase. The left side of $\gamma = 1$, $Z < 1/2$ is the phase in which a bad equilibrium is the super diffusion phase.

When $q = 0.5$, the curves $\gamma = 1$, $Z > 1/2$ in the one peak phase and the two-peaks phase are separated by the curve $\gamma = 0$. When $q \neq 0.5$, the curves $\gamma = 1$, $Z > 1/2$ in the one peak phase and the two-peaks phase are deformed and meet on the curve $\gamma = 0$. Hence, the region surrounded by $\gamma = 1$, $Z > 1/2$ is the super diffusion phase of the good equilibrium.

Finally, we comment on the analog herder case. As discussed in section 2, if we set $\lambda = 2$ and use only the linear terms in (1), we can obtain the analog herder case. In this limit, we can solve (14) and obtain $\bar{v} = 2q - 1$. Equation (14) has only one solution in entire region p. Hence, we do not observe information cascade transition, which is consistent with previous conclusions [13]. On the other hand, in super and normal diffusion phase transitions, $p_{vc1} = 1/2$ does not depend on q, because the gradient of the linear term is constant.

6 Numerical Simulations

To confirm the analytical results, we performed numerical integration of the master equation (3). We perform simulations for the symmetric independent voter case, i.e. $q = 1/2$.

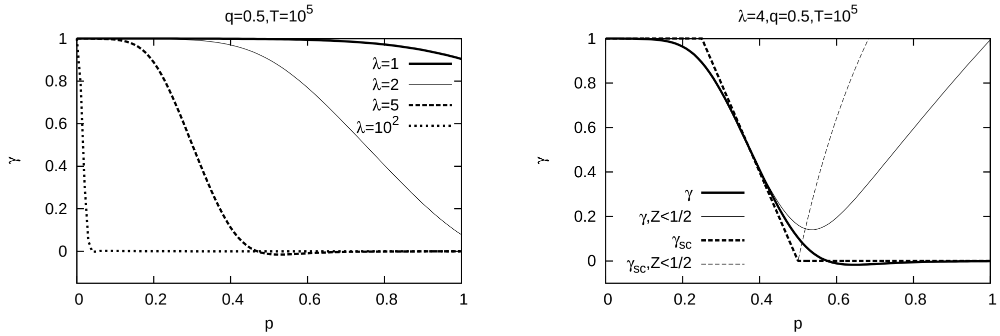
(1) One peak phase: $q = 1/2, \lambda = 4$

Figure 4(a) shows the convergence of the distribution in the one peak phase. We integrated the master equation up to $t = 10^5$. As discussed in previous sections, the critical point of the convergence of this transition is $p_{vc1} \sim 1/\lambda$ and the critical point of information cascade transition is $p_c \sim 2/\lambda$. At the critical point of information cascade transition, the distribution splits into two and the exponent γ becomes 0.[3]

(2) Two-peaks phase: $q = 1/2, \lambda = 4$

Figure 4(b) shows the convergence of the distribution in the two-peaks phase. We consider the case wherein $q = 0.5$ and $\lambda = 4$. In this phase,

[3] Here, we estimate γ from the slope of $\mathrm{Var}(Z(t))$ as $\gamma = \log\{\mathrm{Var}(Z(t-\Delta t))/\mathrm{Var}(Z(t))\}/\log\{t/(t-\Delta t)\}$.



(a) (b)

Figure 4: Convergence of distribution in the one peak phase and the two-peaks phase for the $q = 0.5$ case. The horizontal axis represents the ratio of herders p, and the vertical axis represents the speed of convergence. $\gamma = 1$ is the normal phase, and $0 < \gamma < 1$ is the super diffusion phase. (a) is for the one peak phase. The critical point of the convergence of transition from normal diffusion to super diffusion is $p_{vc1} \sim 1/\lambda$, and the critical point of information cascade transition is $p_c \sim 2/\lambda$. (b) is for the two-peaks phase. The solid lines are obtained by numerical simulations. $\gamma, Z < 0$ is the speed of convergence of the region $Z < 0$, which indicates a bad equilibrium. The dotted line γ_{sc} represents the theoretically obtained curve and $\gamma_{sc}, Z < 0$ represents the theoretically obtained curve for a bad equilibrium. The critical point of information cascade transition is $p_c \sim 0.5$ and the critical point of the convergence of transition is $p_{vc1} \sim 0.25$ and $p_{vc2} \sim 0.675$.

a sequence of voting converges to one of the peaks. We can not determine which peak is chosen in finite time. We divide the distribution into two parts, $Z > 0$ or $Z < 0$, which correspond to two regions around the peaks.

Above the critical point of information cascade transition $p_c \sim 2/\lambda = 0.5$, the speed of the convergence in the region $Z < 0$ is slower than that in a binomial distribution.[4] This indicates the super diffusion phase. In the super diffusion phase, the distance between the two peaks is short and the influence of the other peak reduces the convergence speed. After the transition of convergence to $p_{vc2} = 0.675$, the speed of the convergence is as in a binomial

[4]In the case $q = 0.5$, the convergence in the region $Z > 0$ is identical to that in $Z < 0$ because of the symmetry. We consider only $Z < 0$ here.

distribution $\gamma = 1$. This indicates the normal diffusion phase.

The critical point of convergence, p_{vc2}, in the two-peaks phase is larger than that predicted by theory. We calculated the variance in the region $Z < 0$ using the data from $t = 9.99 \times 10^4$ to $t = 10^5$. The region may be too wide for use in our assumptions, which we described in the previous section (16). To fill the gap between the numerical calculations and theory, we need to perform simulations for a large t.

7 Social Experiments

We conducted simple social experiments for our model. In 2010, we framed 100 questions, each with two choices–knowledge and no knowledge. 31 participants answered these questions sequentially in one group. We performed experiments with two groups. In 2011, we framed 120 questions and 52 participants answered these questions sequentially. We again performed experiments with two groups.

First, they answered the questions without having any information about the others' answers, i.e. their answers were based on their own knowledge. Those who knew the answers selected the correct answers. Hence, we can set $q = 1$. Those who did not know the answers selected the correct answers with a probability of 0.5. Next, the participants were allowed to see all previous participants' answers. Those who did not know the answers referred to this information. We are interested in knowing whether they referred to this information as digital or analog herders.

From the experiments, we can obtain macroscopic behavior q_h [19]. The voters can see all the previous votes, and we could fit the plot by the following functional form:

$$q_h = \frac{1}{2}[a \tanh \lambda \{ \frac{c_1(t)}{(c_0(t) + c_1(t))} - \frac{1}{2} \} + 1]. \tag{20}$$

The difference between (1) and (20) is the constant a. The parameter a denotes the net ratio of the herder that reacts positively to the previous votes. We estimated the parameters $\lambda = 3.80$ and $a = 0.761$ form the experiments.

We can map (20) to (1) as follows:

$$P(k,t) = pq_h + (1-p) = \frac{1}{2}p[a \tanh \lambda \{ \frac{c_1(t)}{(c_0(t) + c_1(t))} - \frac{1}{2} \} + 1] + (1-p)$$

$$= \frac{1}{2}\tilde{p}[\tanh\lambda\{\frac{c_1(t)}{(c_0(t)+c_1(t))} - \frac{1}{2}\} + 1] + (1-\tilde{p})\tilde{q}, \tag{21}$$

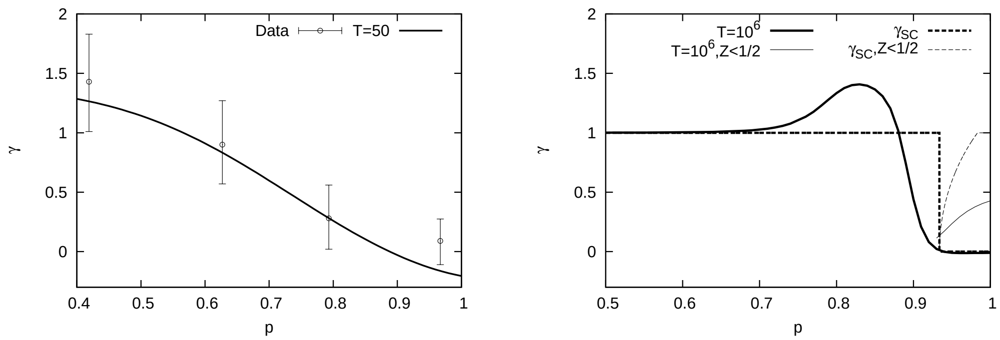
where $\tilde{p} = pa$ and $\tilde{q} = 1 - 1/2 \cdot p(1-a)/(1-pa)$. Using this mapping and conclusions from previous sections, we can theoretically estimate the conclusions for a large limit of t.

Figure 5(a) shows the experimental data and simulation results. $\gamma = 1$ is the normal and $0 < \gamma < 1$ is the super diffusion phase. In the experimental data, when the number of voters is $t = 50$, we can confirm that γ monotonically decreases. For the simulations, we set $t = 50$, and we perform numerical integration of the master equation (3), as described in the previous section. The data points observed in the experiments are on the simulation curve without $p \sim 1$.

Figure 5(b) shows the simulation and theoretical results to study the asymptotic behavior of convergence. For $T = 10^6$, we obtain the simulation curve of the convergence of Z. In $T = 10^6$, represents a simulation curve, γ decreases from 1 to 0 at $p_c \sim 0.9$. This indicates the phase transition from the one peak phase to the two-peaks phase.

Using the conclusions of sections 4 and 5, we can theoretically estimate the conclusions for a large t limit. At $p_c = 0.934$, we observe information cascade transition. γ_{sc} is the theoretical curve that shows information cascade transition at p_c. We observe a difference between the simulation curve $T = 10^6$ and the theoretical curve γ_{sc}. This difference can be reduced by performing a simulation for a large t.

In the one peak phase, below p_c, the peak converges as normal. Above p_c, we observe two peaks. The good equilibrium converges as normal. On the other hand, bad equilibrium converges slower than normal. $\gamma_{sc}, Z < 0$, represents a theoretical curve, shows this phenomenon. At $p_{cv2} = 0.983$, we observe the phase transition of normal and super diffusions, for the bad equilibrium. Above p_{cv2}, both peaks converges as normal. $T = 10^6, Z < 0$ is the simulation curve for the bad equilibrium. We can observe the increase in the convergence speed for the bad equilibrium. The difference between the simulation and the theoretical results has been described in the previous section; this difference may also be reduced by performing a simulation for a large t.



(a) (b)

Figure 5: Behavior of convergence as given by the plot of p vs. γ. The horizontal axis represents the ratio of herders, p, and the vertical axis represents the speed of convergence. $\gamma = 1$ is the normal phase, and $0 < \gamma < 1$ is the super diffusion phase. (a) is for the experimental data and simulation results at $t = 50$. (b) is for simulation results at $t = 10^6$ and theoretical results in the limit $t \rightarrow \infty$. On the basis of the theoretical results, we show the convergence of Z as γ_{sc} and that of $Z < 1/2$ as $\gamma_{sc}, Z < 1/2$. $\gamma_{sc}, Z < 1/2$ shows the theoretical convergence for a bad equilibrium. $T = 10^6$ shows the simulation curve of the convergence of Z. $T = 10^6, Z < 0$ shows the simulation curve for a bad equilibrium. Above the threshold $p_c = 0.934$, we observe in the two-peaks phase.

8 Concluding Remarks

We investigated a voting model that is similar to a Keynesian beauty contest. In the continuous limit, we could obtain stochastic differential equations. The model has two kinds of phase transitions. One is information cascade transition, which is similar to the phase transition of Ising model. In fact, we showed that the stationary condition of our model is same as the equation of state for Ising model. As the herders increased, the model featured a phase transition beyond which a state where most voters make the correct choice coexists with one where most of them are wrong. In this transition, the distribution of votes changed from the one peak phase to the two-peaks phase. These two peaks were the two stable solutions out of the three solutions. In Ising model with an external field, there is only one physical solution out of the three solutions.

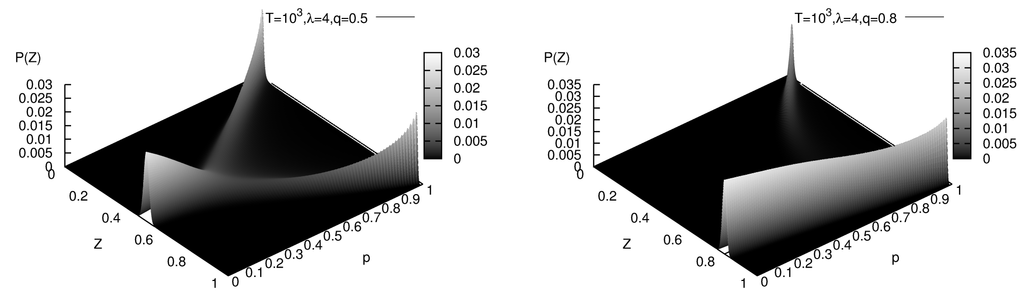
The other transition was the transition of the convergence between super and normal diffusions. In the one peak phase, if herders increased, the variance converged slower than in a binomial distribution. This is the transition from normal diffusion to super diffusion. In the two-peaks phase, the sequential voting converged to one of the two peaks. When $q = 0.5$, this is spontaneous symmetry breaking. In the two-peaks phase, if the herders increased, the variance converged as in a binomial distribution. This is the transition from super diffusion to normal diffusion, which is opposite to the transition in the one peak phase. In other words, the super diffusion phase is sandwiched between normal diffusion phases and is divided by information cascade transition.

We determined the microspecific behavior from social experiments. Using this experimental data, we confirmed the two kinds of phase transitions by performing numerical simulations and conducting analytical studies in the upper t limit.

If the ratio of herders is smaller than p_c, we can distribute correct answers to herders that do not have information by using this voting system [20]. On the other hand, if the ratio of herders is larger than p_c, the system is similar to a Keynesian beauty contest; many voters are herders and there is a case in which more than half the voters make the wrong decision [12]. In the case $p < p_c$, the advantage of this system is evident, and in the case $p > p_c$, the weakness is observed. These are two sides of the same coin and are brought out by information cascade transition.

Figure 6 shows the distributions when (a) $q = 0.5$ and (b) $q = 0.8$. As p increases, the distribution splits into two in both cases. We can think of this as splitting of a particle into two parts by the interactions. In the region near p_c, both the one peak and the two-peaks phases exhibit super diffusion. In this phase, the speed of convergence is slower than that in the normal diffusion phase. We believe that the particles are deformed by the interactions. When $q = 0.5$, a particle is divided continuously. When $q = 0.8$, a second particle appears far from the original particle and is divided discontinuously.

In [10], the authors observed that landslides occurred mostly in countries with a small number of electors in the 2008 US presidential election. This indicates that herders refer to local information, and not global information. In a previous study, we analyzed the case wherein herders could see the r previous votes. This indicates that voters can share information locally. We discussed only information cascade transition and could not discuss transition



(a) (b)

Figure 6: Distributions for (a) $q = 0.5$ and (b) $q = 0.8$, when the number of votes is 10^3. The three axes are Z, p, and the frequency $P(Z)$. As p increases, the distribution splits into two in both cases. In the region near p_c, both the one peak and the two-peaks phases exhibit super diffusion. In this phase, the speed of convergence is slower than that in the normal diffusion phase. We can confirm that the width of the distribution is greater than that in the region far from p_c.

for super and normal diffusions. We consider stochastic differential equations, which are strong tools, and use them for analysis in this study. The case of several candidates remains to be investigated. In this study, we investigated the case of only two candidates[5] In our future studies, we shall consider these cases.

Acknowledgment

This work was supported by Grant-in-Aid for Challenging Exploratory Research, No. 21654054 (SM).

[5] We have previously discussed the cases of several candidates when all voters are analog herders [15]. The probability functions of the share of votes of the candidates obey the gamma distributions.

Appendix A Derivation of model from Bayes' theorem

We estimate the probability that the candidate C_1 is correct by using the previous votes. We assume that the $(t+1)$th voter is a herder. The voter estimates the posterior distribution that C_1 is the correct candidate by using Bayes' theorem. The voter can see t votes.

Herders estimate the probability that C_i, where $i = 0, 1$, is the correct candidate by using the previous votes. $Pr(C_i)$ is the probability that the voter estimates that C_i is the correct candidate. We set the prior distribution as $Pr(C_0) = Pr(C_1) = 1/2$. Here, we assume that the herders believe that all voters are independent and estimate the percentage of correct answers as $\hat{q}$. The posterior distribution is

$$\begin{aligned} Pr(C_1|c_1 = k) &= \frac{1}{2}\frac{t!}{k!(t-k)!}\hat{q}^k(1-\hat{q})^{t-k}, \\ Pr(C_0|c_1 = k) &= \frac{1}{2}\frac{t!}{k!(t-k)!}(1-\hat{q})^k\hat{q}^{t-k}, \end{aligned} \tag{22}$$

where $c_1 = k$ indicates that the number of votes for C_1 is k before the voter votes. We can obtain

$$\frac{Pr(C_1|c_1 = k)}{Pr(C_0|c_1 = k)} = (\frac{\hat{q}}{1-\hat{q}})^{2k-t} = e^{2\lambda(k-\frac{t}{2})/t}, \tag{23}$$

where $\lambda = t\log\frac{\hat{q}}{1-\hat{q}}$. Here, λ increases as t increases. Hence, the herder can calculate the probability that the candidate C_1 is correct when the number of votes for C_1 is k:

$$Pr(C_1|c_1 = k) = \frac{1}{2}[\tanh\lambda(\frac{k}{t} - \frac{1}{2}) + 1]. \tag{24}$$

In the region $t >> 1$, $\lambda \rightarrow \infty$, a voter believes the public perception, and the behavior of herders becomes similar to that of digital herders without $\hat{q} = 1/2$. When $\hat{q} = 1/2$, the herders estimate that the previous votes are not useful to estimate the correct answer and $\lambda = 0$.

Appendix B Derivation of stochastic differential equation

We use $\delta X_\tau = X_{\tau+\epsilon} - X_\tau$ and ζ_τ, a standard i.i.d. Gaussian sequence; our objective is to identify the drift f_τ and the variance g_τ^2 such that

$$\delta X_\tau = f_\tau(X_\tau)\epsilon + \sqrt{\epsilon} g_\tau(X_\tau)\zeta_{\tau+\epsilon}. \tag{25}$$

Given $X_\tau = x$, using the transition probabilities of Δ_n, we get

$$\begin{aligned} \mathrm{E}(\delta X_\tau) &= \epsilon \mathrm{E}(\Delta_{[\tau/\epsilon]+1} - \Delta_{[\tau/\epsilon]}) = \epsilon(2p_{[\frac{l/\epsilon+\tau/\epsilon}{2}],\tau/\epsilon} - 1) \\ &= \epsilon[(1-p)(2q-1) + p\tanh(\frac{\lambda x}{2\tau})]. \end{aligned} \tag{26}$$

Then, the drift term is $f_\tau(x) = (1-p)(2q-1) + p\tanh(\lambda x/2\tau)$. Moreover,

$$\sigma^2(\delta X_\tau) = \epsilon^2[1^2 p_{[\frac{l/\epsilon+\tau/\epsilon}{2}],\tau/\epsilon} + (-1)^2(1 - p_{[\frac{l/\epsilon+\tau/\epsilon}{2}],\tau/\epsilon})] = \epsilon^2, \tag{27}$$

such that $g_{\epsilon,\tau}(x) = \sqrt{\epsilon}$. We can obtain X_τ such that it obeys a diffusion equation with small additive noise:

$$\mathrm{d}X_\tau = [(1-p)(2q-1) + p\tanh(\frac{\lambda X_\tau}{2\tau})]\mathrm{d}\tau + \sqrt{\epsilon}. \tag{28}$$

Appendix C Behavior of solutions of stochastic differential equation

We consider the stochastic differential equation

$$\mathrm{d}x_\tau = (\frac{l x_\tau}{\tau})\mathrm{d}\tau + \sqrt{\epsilon}, \tag{29}$$

where $\tau \geq 1$. If we set $l = p\lambda(1-\bar{v}^2)/2$, (29) is identical to (18). Let σ_1^2 be the variance of x_1. If x_1 is Gaussian ($x_1 \sim \mathrm{N}(x_1, \sigma_1^2)$) or deterministic ($x_1 \sim \delta_{x1}$), the law of x_τ ensures that the Gaussian is in accordance with density

$$p_\tau(x) \sim \frac{1}{\sqrt{2\pi}\sigma_\tau} \mathrm{e}^{-(x-\mu_\tau)^2/2\sigma_\tau^2}, \tag{30}$$

where $\mu_\tau = \mathrm{E}(x_\tau)$ is the expected value of x_τ and $\sigma_\tau^2 \equiv \nu_\tau$ is its variance. If $\Phi_\tau(\xi) = \log(\mathrm{e}^{\mathrm{i}\xi x_\tau})$ is the logarithm of the characteristic function of the law of x_τ, we have

$$\partial_\tau \Phi_\tau(\xi) = \frac{l}{\tau}\xi\partial_\xi \Phi_\tau(\xi) - \frac{\epsilon}{2}\xi^2, \tag{31}$$

and

$$\Phi_\tau(\xi) = \mathrm{i}\xi\mu_\tau - \frac{\xi^2}{2}\nu_\tau. \tag{32}$$

Identifying the real and imaginary parts of $\Phi_\tau(\xi)$, we obtain the dynamics of μ_τ as

$$\dot{\mu}_\tau = \frac{l}{\tau}\mu_\tau. \tag{33}$$

The solution for μ_τ is

$$\mu_\tau = x_1\tau^l. \tag{34}$$

The dynamics of ν_τ are given by the Riccati equation

$$\dot{\nu}_\tau = \frac{2l}{\tau}\nu_\tau + \epsilon. \tag{35}$$

If $\nu \neq 1/2$, we get

$$\nu_\tau = \nu_1\tau^{2l} + \frac{\epsilon}{1-2l}(\tau - \tau^{2l}). \tag{36}$$

If $l = 1/2$, we get

$$\nu_\tau = \nu_1\tau + \epsilon\tau\log\tau. \tag{37}$$

We can summarize the temporal behavior of the variance as

$$\nu_\tau \sim \frac{\epsilon}{1-2l}\tau \qquad \text{if} \quad l < \frac{1}{2}, \tag{38}$$

$$\nu_\tau \sim (\nu_1 + \frac{\epsilon}{2l-1})\tau^{2l} \qquad \text{if} \quad l > \frac{1}{2}, \tag{39}$$

$$\nu_\tau \sim \epsilon\tau\log(\tau) \qquad \text{if} \quad l = \frac{1}{2}. \tag{40}$$

This model has three phases. If $l > 1/2$ or $l = 1/2$, x_τ/τ converges slower than in a binomial distribution. These phases are the super diffusion phases. If $0 < p < 1/2$, x_τ/τ converges as in a binomial distribution. This is the normal phase [13].

References

[1] Tarde G 1890 *Les lois de l'imitation*

[2] Milgram S, Bickman L, and Berkowitz L 1969 *J. Per. Soc. Psyco.* **13** 79

[3] Partridge B L 1982 *Sci. Am.* **245** 90

[4] Couzin I D, Krause J, James R, Ruxton G R, and Franks N R 2002 *J. Theor. Biol.* **218** 1

[5] Galam G 1990 *Stat. Phys.* **61** 943

[6] Cont R and Bouchaud J 2000 *Macroecon. Dyn.* **4** 170

[7] Eguíluz V and Zimmermann M 2003 *Phys. Rev. Lett.* **85** 5659

[8] Stauffer D 2002 *Adv.Complex Syst.* **7** 55

[9] Curty P and Marsili M 2006 *JSTAT* P03013

[10] Araújo N A M, Andrade Jr. J S, and Herrmann H J 2010 *PLOS ONE* **5** e12446

[11] Bikhchandani S, Hirshleifer D, and Welch I 1992 *J. Polit. Econ.* **100** 992

[12] Keynes J M 1936 *General Theory of Employment Interest and Money* (Macmillan Cambridge University Press)

[13] Hisakado M and Mori S 2010 *J. Phys. A* **43** 3152 7

[14] Mori S and M Hisakado arXiv:1006.4884 Power law convergence of win bet fraction and component ratio of herding better in a racetrack betting market Preprint

[15] Mori S and Hisakado M 2010 *J. Phys. Soc. Jpn.* **79** 034001

[16] Hisakado M, Kitsukawa K, and Mori S 2006 *J. Phys. A* **39** 15365

[17] Hod S and Keshet U 2004 *Phys. Rev. E* **70** 11006

[18] Hisakado M and Mori S 2011 *J. Phys. A* **22** 275204

[19] Mori S, Hisakado M, and Takahashi T 2011 *Phase transition to two-peaks phase in an information cascade voting experiment*

[20] Hayek F A 1962 *Rules, Perception and Intelligibility* (Oxford University Press)